UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/608,526	06/30/2000	Shuo DI	MSI-449US	1121

22801 7590 02/19/2004

LEE & HAYES PLLC
421 W RIVERSIDE AVENUE SUITE 500
SPOKANE, WA 99201

EXAMINER	
STEVENS, THOMAS H	

ART UNIT	PAPER NUMBER
2123	5

DATE MAILED: 02/19/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/608,526	DI ET AL.	
	Examiner	Art Unit	
	Thomas H. Stevens	2123	

-- Th MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *30 June 2000 and 03 October 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-29* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-29* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *30 June 2000* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Abstract

1. Applicant is reminded of the proper language and format for an abstract of the disclosure.

The abstract should be in narrative form and generally limited to a single paragraph on a separate sheet within the range of 50 to 150 words. In this case, the abstract is less than 50 words.

Drawings

2. Figures 1, 4 and 5 should be labeled as prior art.

Claim Objections

3. It is notified that claims 1 and 24 do not recite preambles, which are related to the limitations.

Claim Rejections - 35 USC § 112

4. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter, which the applicant regards as his invention.

5. Claim 4 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The phrase "the minimum threshold of 3" is vague and indefinite.

6. Claims 25,27-29 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Each of these stated claims discloses the theme of "to improve language predictive capability" or "improve one or more language model performance attributes", which is mere allegation of benefits or advantages, but does not further limit the claims.

Claim Rejections - 35 USC § 101

7. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

8. Claims 1-29 are rejected under 35 U.S.C. 101 because the claimed invention is directed to a mathematical algorithm. The examiner respectfully submits that the applicants have not claimed a practical application. An invention which is eligible for patenting under 35 U.S.C. § 101 is in the "useful arts" when it is a machine, manufacture, process or composition of matter, which produces a concrete, tangible, and useful result.

The examiner respectfully submits, under current PTO practice, that the claimed invention does not recite a tangible or concrete result. The claims are not tangible because they appear to recite, in the narrower claims, a mathematical algorithm, namely reviewing and correcting data language or structures within a statistical model, which is confined or limited space that doesn't have specific preprocessing or post solution activity.

Claim Rejections - 35 USC § 102

9. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

10. Claims 1-29 are rejected under 35 U.S.C. 102(b) as being anticipated by Kukich (Paper: 1992). Kukich's paper teaches methods of correcting words and non-words; non-word error detection; isolated-word error correction; and context-dependent word correction through statistical analysis(abstract).

Claim 1: A method comprising: assigning each of a plurality of segments comprising a received corpus to a node (pg. 403, column 1, paragraph 2, lines17-32 and pg. 380, column 2 lines 17-20) in a data structure denoting dependencies between nodes; and calculating a transitional probability (pg. 401, column 1, section 2.2.5, lines 5-7) between each of the nodes in the data structure.

Claim 2: A method according to claim 1, further comprising: calculating a frequency (pg. 413, title section 3.1) for each elemental item of the segment; and removing nodes of the data structure associated with items which do not meet a minimum frequency threshold (pg. 426, column 1, lines 4-14).

Claim 3: A method according to claim 2, wherein the frequency of the item is calculated by counting item occurrences throughout the subset and/or corpus (pg. 406, column 1, lines 1-14).

Claim 4: A method according to claim 2, wherein the minimum threshold (pg. 406, column 2, paragraph 3) is three.

Claim 5: A method according to claim 1, further comprising: managing storage of the data structure across a system memory of a computer system (pg. 381,columns 1 and 2, last three lines and first three lines, respectively) and an extended memory of the computer system.

Claim 6: A method according to claim 5, wherein the step of managing storage of the data structure comprises: identifying least recently used nodes of the data (pg. 428, column 2, lines 20-25) structure; and storing the least recently used nodes of the data structure in the extended memory of the computer system when the data structure is too large to store completely within the system memory (pg. 381,columns 1 and 2, last three lines and first three lines, respectively).

Claim 7: A method according to claim 5, wherein the extended memory of the computer system comprises one or more files on an accessible mass storage device (pg. 381,columns 1 and 2, last three lines and first three lines, respectively).

Claim 8: A method according to claim 7, wherein the data structure represents to a language model, spread across one or more elements of a computing system memory subsystem (pg. 403, column 1, paragraph 2, lines17-32).

Claim 9: A method according to claim 1, wherein calculating a transition probability includes calculating a Markov transitional probability (pg. 401, columns 1 and 2, lines 12-13 and 1, respectively) between nodes.

Claim 10: A storage medium comprising a plurality of executable instructions including at least a subset of which that, when executed by a processor, implement a method according to claim 1 (pg. 403, column 1, paragraph 2, lines17-32).

Claim 11: A method for predicting a likelihood of an item in a corpus(pg. 403, column 1, paragraph 2, lines17-32) comprised of a plurality of items, the method comprising: building a data structure of corpus segments representing a dynamic context of item dependencies within the segments; calculating the likelihood of each item based, at least in part, on a likelihood of preceding items within the

dynamic context; and iteratively re-segmenting the corpus to improve the calculated likelihood of item dependencies (pg. 432, column 1, paragraph 2).

Claim 12: A method according to claim 11, wherein the method of building a dynamic context of preceding dependent items comprises: analyzing the data structure representing the language model; identifying all items with dependencies to or from the item; and a using all items with dependencies to or from the item as the dynamic context (pg. 432).

Claim 13: A method according to claim 11, wherein the language model includes frequency information for each item within the model (pg. 381, column 2, paragraph 3).

Claim 14: A method according to claim 13, wherein calculating the likelihood of the item comprises: calculating a Markov (pg. 401, columns 1 and 2, lines 12-13 and 1, respectively) transition probability for the item based, at least in part, on the frequency (pg. 381, column 2, paragraph 3) of the items comprising the dynamic context.

Claim 15: A method according to claim 11, wherein calculating the likelihood of the item comprises: calculating a Markov (pg. 401, columns 1 and 2, lines 12-13 and 1, respectively) transition probability for the item given the dynamic context of items.

Claim 16: A storage medium having stored thereon a plurality of executable instructions including instructions which, when executed by a host computer implement a method according to claim 11 (pg. 432, column 1, paragraph 2).

Claim 17: A data structure, generated by a computer system as a statistical language model, the data structure comprising: one or more root nodes (pg. 383, column 1, paragraph 3, lines 6-8); and a plurality of subordinate nodes, ultimately linked to a root node, cumulatively comprising one or more sub-trees, wherein each node of a to sub-tree represents (pg. 385, column 1, paragraph 1, lines 4-6) one or more items of a corpus and includes a measure of a Markov transition (pg. 401, column 1, paragraph 3) probability between the node and another linked node.

Claim 18: A data structure according to claim 17, wherein the root node represents a common root item for all subordinate nodes in the one or more sub trees (pg. 402, column 2, paragraph 3, lines 10-12).

Claim19: A data structure according to claim 17, wherein the Markov (pg. 401, column 1, paragraph 3 and column 2, paragraph 1) is transition probability is a measure of the likelihood of a transition from one node to another node based, at least in part, on the one or more items represented by each of the nodes (pg. 403, column 1, lines 17-26).

Claim 20: A data structure according to claim 17, wherein the items include one or more of a character, a letter, a number, and combinations thereof (pg. 403, column 2, lines 6-14).

Claim 21: A data structure according to claim 17, wherein the data structure represents a dynamic order Markov (pg. 401, column 1, paragraph 3 and column 2, paragraph 1) model (DOMM) language model of the textual source.

Claim 22: A storage medium comprising a plurality of executable instructions which, when executed by a processor, implement a data structure according to claim 17 (pg. 432; and 382, column 1, last sentence with column 2, first paragraph).

Claim 23: A memory subsystem in a computer system including one or more of a cache memory (pg. 383, column 1, paragraph 2, lines 4-8), a system memory and extended memory having information stored therein which, when interpreted by a processor of the computer system (pg. 381, column 1, paragraph 3, lines 9-10; and column 2, lines 1-3) represent a data structure according to claim 17.

Claim 24: A modeling agent comprising: a controller, to receive a corpus (pg. 403, column 1, paragraph 2, lines17-32 and pg. 380, column 2 lines 17-20); and a data structure generator, responsive to and selectively invoked by the controller, to assign each of a plurality of segments comprising the received

corpus to a node in a data structure denoting dependencies between nodes; wherein the modeling agent calculates a transitional probability (pg. 401, column 1, section 2.2.5, lines 5-7) between each of the nodes of the data structure to determine a predictive capability of a language model represented by the data structure and iteratively re-segments (pg. 432, column 2, lines 1-16) the received corpus until a threshold (pg. 426, column 1, lines 4-14) predictive capability is reached .

Claim 25: A modeling agent according to claim 24, the data structure generator comprising: a dynamic segmentation function, to iteratively re-segment the received corpus to improve language model predictive capability (pg. 432).

Claim 26: A modeling agent according to claim 24, the data structure generator comprising: a frequency analysis function (pg. 406, column 1, lines 1-14) to analyze a frequency of occurrence of segments within the corpus.

Claim 27: A modeling agent according to claim 26, wherein segments that do not meet a frequency (pg. 406, column 1, lines 1-14) of occurrence threshold (pg. 426, column 1, lines 4-14) are removed from the data structure, reducing data structure size and improving language model predictive capability.

Claim 28: A storage medium comprising a plurality of executable instructions including at least a subset of which, when executed, implement a language modeling agent to assign each of a plurality of segments of a received corpus to a nodes pg. 403, column 1, paragraph 2, lines17-32 and pg. 380, column 2 lines 17-20) in a data structure denoting dependencies between nodes, and to calculate a transitional probability between each of the nodes in the data structure to determine a predictive capability of a language model denoted by the data structure, wherein the modeling agent dynamically re-segments the received corpus to remove segments (pg. 432, column 2, lines 1-16) which do not meet a minimum frequency threshold (pg. 406, column 2, paragraph 3) to improve one or more language model performance attributes.

Claim 29: A storage medium according to claim 28, wherein the one or more language model performance attributes include a predictive capability (pg. 431, figure 1).

Correspondence Information

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tom Stevens whose telephone number is (703) 305-0365, Monday-Friday (8:30 am- 5:30 pm) or contact Supervisor Mr. Kevin Teska at (703) 305-9704. The fax number for the group is 703-872-9306.

Any inquires of general nature or relating to the status of this application should be directed to the Group receptionist whose phone number is (703) 305-3900.

February 11, 2004

THS
THS

HUGH JONES Ph.D.
PRIMARY PATENT EXAMINER
TECHNOLOGY CENTER 2100